DOCUMENT RESUME

ED 194 566 TM 800 695

AUTHOR McKee, Barbara G.
TITLE Student's Course-Oriented Attitude Change and Student Ratings of Instruction: A Canonical Variate Analysis.
PUB DATE Apr 80
NOTE 36p.: Paper presented at the Annual Meeting of the American Educational Research Association (64th, Boston, MA, April 7-11, 1980).

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Attitude Measures: *Course Evaluation: Higher Education: *Multivariate Analysis: *Predictive Measurement: Predictor Variables: Pretests Posttests: Rating Scales: *Student Attitudes: *Student Evaluation of Teacher Performance: Test Reliability
IDENTIFIERS Adjective Rating Scale: Instructional Development Effective Assessment

ABSTRACT

Canonical variate analysis was employed to measure the direction and intensity of relationships between student ratings of instruction and students' course-oriented attitude change. Change was defined as the residual post-course attitude scores when pre-course attitude scores served as predictors. The Adjective Rating Scale (ARS), a measure of course-oriented attitudes, was administered to students in 62 college classes on the first day of classes and during the final two weeks of the semester. ARS factors were labeled interest/affect, practical value, negative affect, and difficulty. The Instructional Development and Effectiveness Assessment (IDEA) Survey, a student rating of the quality of instruction, was administered to the students at the end of the course only. Two dimensions (course difficulty and positive interest/affect) were significantly related to residualized attitude toward the quality of a course. Students post-course attitudes that differed significantly from what was predicted were related not only to course variables, but also to the students' perceptions of instructor behaviors. Little information was available on either the test-retest reliability or the parallel form reliability of the measures used. (Author/RL)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED194566

U S DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

Student's Course-Oriented Attitude Change and Student Ratings of Instruction: A Canonical Variate Analysis

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
B. McKee

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Barbara G. McKee

Assistant Professor and Research Associate

Department of Educational Research & Development

Rochester Institute of Technology

Rochester, New York 14623

Presented at the Annual Meeting

of the American Educational Research Association

Boston, Massachusetts

April, 1980

TM 800 559

Abstract

This study employed canonical variate analysis with a redundancy measure to examine the direction and intensity of relationships between student ratings of instruction and students' course-oriented attitude change. Change was defined as the residual post-course attitude scores when pre-course attitude scores served as predictors. The measures were administered to sixty college classes on the first and last days of class. Results indicated two strong relationships between the attitude residuals and rating scores. Educational researchers interested in student ratings of instruction were urged to integrate their work with social science literature on the structure, acquisition, and change of attitudes.

Students' Course - Oriented Attitude Change and Student Ratings of Instruction: A Canonical Variate Analysis

There currently is considerable interest among college students, faculty, and administrators in evaluating faculty teaching effectiveness. "Included in this demand for evaluation has been the widespread use of formal student appraisal of courses and instructors" (Costin, Greenough, & Menges, 1971, p. 511). Reasons given for this surge of interest in student ratings of instruction range from pressures for accountability to student unrest (French-Lazovik, 1972).

Although the actual number of institutions that currently use formal student ratings of instruction is uncertain (Holzemer, 1975), the desire of faculty to see formal student ratings of instruction used as evidence of teaching effectiveness has been demonstrated (Scott, 1975). Menges (1973) stated that the teacher and his/her students are the only direct sources of information on what occurs in the college classroom, and, if classroom events are one focus for the evaluation of teaching, observations by students seem essential. In an extensive review of previous research, Costin, Greenough, and Menges (1971) concluded that student ratings of instruction are generally found to be stable, internally reliable, and valid with respect to many criteria of teaching effectiveness.

Other literature reviews and research projects have taken exception to the findings of Costin, Greenough, and Menges. Follman (1975) concluded that the personality characteristics of college student raters influenced

their ratings of their instructors' teaching effectiveness, and Hogan (1973) found that the course setting affects student ratings of instruction. An unsettling study by Nauftulin, Ware, and Donnelly (1973) has received a great deal of attention. These authors asked an actor to deliver a "charismatic but nonsubstantive" lecture to a group of professional mental health workers (including psychiatrists, psychologists, educators, and administrators). Despite the lecture's nonsubstantive character, the audience rated it and the lecturer very favorably. The debate over the utility of student ratings of instruction was recently summed up by James Kulik: "Student ratings may be irrelevant and misleading, or they may be useful, convenient, reliable, and valid" (1974, p. 51).

Since student ratings of instruction have become a part of the evaluation of instruction in many colleges, it is desirable to know as much as possible about the properties of such ratings. The literature on college students' ratings of instruction dates from the 1920's, but there is little consistent agreement as to what ratings actually measure (Centra, 1972).

One of the unexplored areas of student ratings of instruction is the relationship between such ratings and students' course-oriented attitudes. Rosenshine recently noted the confusion between the terms "student ratings of" and "student attitudes toward" instruction.

> . . . the term student ratings has been used in this discussion section, but the term student attitudes was used in earlier paragraphs. The items used by Flanders to measure student attitude seem quite similar to the items in the student rating form used by Gage (1972). Currently, student ratings of a teacher and student attitudes toward a teacher are terms with overlapping meaning, but it may be necessary to begin to separate these terms in future research.
> (Rosenshine, 1973, p. 180)

The relationship between attitudes and ratings has been explored neither in the literature on student ratings of instruction nor in the social science literature on attitude measurement.

There is a clear paucity of studies that attempt to differentiate between student attitudes toward a course and student ratings of a course. If a distinction is to be made between these two terms, further studies must be conducted. Researchers who have investigated student ratings have, for the most part, simply ignored any distinction between these two terms and have used attitudes toward instruction and ratings of instruction interchangeably if they used the term attitude at all. Psychologists and other social scientists interested in the structure, acquisition, and change of attitudes, on the other hand, have not traditionally been concerned with attitudes toward an object as limited in time and setting as a specific course.

As Rokeach (1970) has pointed out, attitude theorists have generally been more interested in the theory and measurement of attitudes toward objects, across situations, than in the theory and measurement of attitudes toward situations, across objects (p. 118). He goes on to say that we have, for example, scales that measure attitudes toward the Negro, the church, labor, and socialism. We do not have scales that measure attitudes toward such situations as managing or eating in a restaurant, being a passenger or driver of a bus, buying or selling real estate. As a result, the study of attitudes-toward-situations has become more or less split off from the study of attitudes-toward-objects. To account for the ways people behave in specific situations, personality psychologists typically use terms or concepts such as traits, and social psychologists use terms or concepts such as role.

"The splitting-off of attitude-toward-situation from attitude-toward-objects has, in the writer's (Rokeach) opinion, severely retarded the growth of attitude theory" (p. 119). It has resulted in a failure to appreciate that an attitude object is always encountered within some situation, about which we also have an organized attitude. Rokeach goes on to say that this splitting-off has resulted in unsophisticated attempts to predict behavior accurately on the basis of a single attitude-toward-object, ignoring the equally relevant attitude-toward-situation. This preoccupation with attitude objects that occurs in a variety of settings partially explains why attitude theorists have not traditionally been concerned with an object as specific as a course. A course is encountered in (in many ways, is) a situation, whereas attitudes toward education or attitudes toward an "ideal" professor can be considered situationless.

The relationship between student ratings of instruction and students' course-oriented attitudes should be further explored. Such explorations would constitute one method of shedding light on the still murky problem of what student ratings of instruction actually measure.

Methodology

Sample

Sixty-two faculty members at one of the State University of New York Colleges participated in this study. The faculty members were chosen on several pragmatic criteria. An attempt was made to include every department in the college, and to include some faculty who traditionally teach upper division courses and some who traditionally teach lower division courses in order to force a larger representation of students, and finally, the faculty involved (with a few exceptions) were known to the principal investigator.

The 62 classes in this study reported an average enrollment of 34.8 students per class. This compares to an average class size of all courses offered at the college in the fall of 1976 of 26.5 students. The average attendance per class for the completion of the post-course instruments equaled 26.7 students per class. This was an average completion rate of 76%.

Chi-Square comparisons revealed that the faculty participating in this study did not differ (at the .01 level of significance) from the total faculty population at the college on the demographic variables of sex, rank, years of higher education experience, department, and division. The sample included a slightly smaller percent of faculty having already received tenure compared to the total faculty population. This difference was significant at the .01 level but not at the .05 level of significance.

The demographic variables of sex, age, full or part-time, and class were available both for students who completed the instruments used in this study and for the total undergraduate population of students at the college. With one exception, no significant differences emerged between the two groups. The sample population contained slightly more (56%) juniors and seniors than did the total undergraduate student population (50%).

Instrumentation

The Adjective Rating Scale (ARS), a measure of course-oriented attitudes, was administered to the 62 classes on the first day of classes and during the final two weeks of the 1976 fall semester. The Instructional Development and Effectiveness Assessment (IDEA) Survey, a student rating of instruction survey, was administered to the students at the end of the course only.

Adjective Rating Scale. The Adjective Rating Scale (ARS) is a measure of course and program-oriented attitude that has been developed by E.F. Kelly at the Center for Instructional Development, Syracuse University (Kelly, 1975). The scale consists of twenty-four adjectives which are rated against the stimulus, "I have found (or expect) this course to be:" The ratings are taken over a four-point scale, "Extremely," "Very," "Somewhat," "Not at All."

The twenty-four terms on the instrument originally were selected from a large set of adjectives used by students at Syracuse University when they were asked to respond to the stimulus item, "List the three words that best describe the course you have just completed." Terms are chosen for their common usage and bipolarity (Kelly et al., 1976). (See Appendix for copy of instrument).

To investigate the instrument, Kelly and Greco (Kelly et al., 1976) analyzed the ratings of precourse expectations from 769 respondents across six courses.[1] Their study involved two different computer programs and a variety of factor analytic techniques.

While factor structures appeared somewhat different from group to group, their results suggest "a definite discernible and stable factor structure which accounts for between 60 and 70 percent of the variance." These factors were labeled as Practical Value, Emotional Appeal, Dullness, Interest Value, and Difficulty. Kelly (1975) reported Alpha coefficients ranging from .71 to .85 for the five factors.

[1]The six courses ranged from freshman to graduate level and included English 101, Psychology 205, Marketing 355, General Education 105, Mathematics 117, and several Management courses.

Instructional Development and Effectiveness Assessment (IDEA). The Instructional Development and Effectiveness Assessment Survey was developed by Hoyt (1969a) at Kansas State University. The survey is now copyrighted by the Center for Faculty Evaluation and Development in Higher Education at Kansas State.

The survey was originally developed on the assumption that one method of defining teaching effectiveness was the students' self-ratings of their progress on specific course objectives. Such objectives should represent those that the faculty member felt were emphasized in the course being rated.

The IDEA Survey Form used in this study (see Appendix) is the third version of the form developed by Hoyt. This final version was based on an administration of the instrument in 1973 and 1974 to over 1400 classes at Kansas State University.

Analysis and Results

The analysis for this study was guided by the attempt to answer the following question:

> What is the relationship between students' course-oriented attitude change and student ratings of instruction?

The question was analyzed in several steps. First, students who had completed both the pre-course ARS and the post-course ARS and IDEA Survey were matched by student I.D. numbers. Since this required that the student have completed both the pre and post instruments, and that he or she have correctly included his/her student I.D. number on both the instruments, the total number of students and classes was reduced. (Some of the smaller classes collapsed to only 1 or 2 students and were dropped). The matched pre-post group consisted of 1223 students in 57 classes.

The pre and post ARS scores were factor analyzed using a principal components solution with varimax rotation. The result of the factor analysis of the ARS are reported in Table 1.

The IDEA was also factor analyzed using the matched subset of students. An eight-factor structure was obtained. The results of the factor analysis of the IDEA Survey are summarized in Table 2.

Two sets of factor scores based on the merged pre-post factor analysis of the ARS were computed for each student, one set for the pre-course ARS and one for the post-course ARS. The pre and post ARS factor scores for each student were submitted to the SPSS subprogram Regression using a simple regression design. Residuals for each of the four post-course factors were computed for each student. Tables 3 and 4 summarize the results of the multiple regression analysis.

Class means were then calculated for each of the four ARS residual scores and IDEA factor scores. The two sets of variables, class means of ARS residuals and class means of IDEA factor scores were submitted to the Cooley and Lohnes canonical correlation program.

This study can be thought of as an extension of an early report which examined the relationship between the post-course ARS factors and the IDEA Survey factors through the use of canonical variate analysis (see McKee, 1978). This study examined the relationship between pre and post-course attitude change as measured by the ARS residuals and the IDEA factors.

Overall means, standard deviations and interfactor correlations of both sets of measures are reported in Table 5.

Table 1

FOUR FACTOR STRUCTURE OF PRINCIPAL COMPONENTS ANALYSIS WITH VARIMAX ROTATION FOR COMBINED PRE-AND POST-COURSE ADJECTIVE RATING SCALE

	Factors				
Variable	Interest/ Affect	Practical Value	Negative Affect	Difficulty	Communality
1. Interesting	.75	.23	.28	.00	.63
2. Boring	-.51	-.03	.61	.01	.63
3. Relevant	.24	.73	-.21	-.01	.63
4. Informative	.41	.50	-.20	.17	.49
5. Difficult	-.10	.01	.11	.84	.73
6. Good	.66	.32	-.27	.10	.62
7. Stimulating	.73	.24	-.26	.06	.66
8. Irrelevant	-.04	-.42	.63	.05	.58
9. Worthwhile	.47	.66	-.22	-.05	.71
10. Valuable	.37	.75	-.17	.06	.73
11. Necessary	.06	.71	-.09	.14	.53
12. Dull	-.48	-.07	.66	.00	.67
13. Challenging	.22	.13	-.08	.78	.68
14. A Waste	-.14	-.27	.75	.00	.66
15. Practical	.25	.72	-.12	-.06	.59
16. Demanding	.00	.10	.02	.86	.75
17. Different	.44	-.12	.18	.31	.33
18. Enjoyable	.77	.19	-.19	.16	.70
19. Enlightening	.67	.32	-.15	.05	.58
20. Exciting	.77	.21	-.12	.03	.65
21. Rewarding	.58	.44	-.13	.01	.55
22. Provocative	.60	.15	.05	.13	.39
23. General	-.03	.12	.34	-.34	.25
24. Useless	-.10	-.32	.67	.05	.56
Eigenvalue	8.68	2.42	1.77	1.46	
% of Variance	36.2	10.1	7.4	6.1	
Cum. % of Variance	36.2	46.3	53.7	59.8	

Table 2

EIGHT FACTOR STRUCTURE OF PRINCIPAL COMPONENTS ANALYSIS WITH VARIMAX ROTATION OF IDEA SURVEY MATCHED GROUP (N=1223)

Variable*	Organization	Faculty Empathy for Students	Student Cognitive Gain	Examinations (Tests)	Student Personal Gain	Teaching Style	Student Workload (non-reading)	Student Workload (Reading) Assignments	Communality
1	-.04	.78	.11	-.02	.09	.20	-.10	-.02	.67
2	.10	.68	.19	-.08	.14	.25	.08	-.22	.67
3	.10	.80	.09	-.08	.12	.14	-.04	.02	.67
4	.16	.17	.09	-.11	.06	.64	.10	.13	.52
5	.21	.40	.20	-.05	.15	.44	.09	-.09	.49
6	-.01	-.14	-.12	.77	-.02	-.07	.00	-.01	.64
7	.22	.12	-.02	-.06	.16	.74	.08	.05	.65
8	.43	.15	.30	-.04	.06	.39	.05	.21	.50
9	-.17	-.12	-.24	.11	-.14	-.69	.01	.00	.61
10	.71	.10	.22	-.01	.06	.16	-.04	.10	.61
11	.44	.30	-.05	-.19	.32	.03	.09	.06	.44
12	-.25	-.03	-.02	.73	-.10	-.11	.09	.07	.63
13	.24	.64	-.02	-.19	.12	.00	-.06	.13	.55
14	.59	.13	.23	-.20	.09	.24	.04	-.12	.54
15	.14	.24	.29	.04	.23	.40	.37	.15	.53
16	.74	.09	.15	-.05	.05	.15	.04	.04	.62
17	.63	.03	.26	-.15	.06	.37	.02	-.10	.64
18	.24	.27	.27	.05	.09	.24	-.32	.48	.61
19	-.10	-.11	-.11	.79	-.11	-.10	.15	.11	.72
20	.23	.30	.33	.03	.22	.51	-.02	.24	.63
21	.26	-.03	.73	.04	-.06	.13	.19	.03	.66
22	.20	.09	.74	-.07	.09	.09	.09	.04	.63
23	.10	.28	.65	-.04	.25	.05	.02	.08	.58
24	.18	.15	.68	-.11	.20	.02	.13	-.05	.59
25	.04	.00	.60	-.06	.11	.08	.19	.04	.43
26	.06	.19	.38	.01	.64	.18	.08	-.10	.65
27	.15	.13	.36	-.05	.56	.07	.21	.07	.55
28	.04	.00	.09	-.12	.66	.29	-.04	.02	.54
29	.11	.22	.03	-.10	.75	.02	-.04	-.06	.63
30	.07	.09	.39	.00	.61	.21	-.15	.06	.61
31	-.03	-.09	.03	.12	-.05	.07	.22	.79	.70
32	.04	.02	.09	.03	.07	.04	.72	-.20	.57
33	-.04	-.09	.05	.15	-.14	.13	.70	.30	.65
34	.40	-.01	.43	-.05	.10	.30	.10	-.05	.43
35	.14	-.13	.26	.12	.04	.17	.66	.29	.66
36	.00	-.13	.45	-.06	.15	.26	-.20	.08	.36
37	.29	.10	.43	-.15	.14	.55	-.01	-.09	.63
38	.18	.05	.62	-.08	.21	.34	-.18	.13	.60
*See Preliminary Analysis for IDEA items.									
Eigenvalue	10.396	3.053	1.993	1.959	1.522	1.358	1.134	1.013	
% of Variance	27.4	8.0	5.2	5.2	4.0	3.6	3.0	2.7	
Cumulative % of Variance	27.4	35.4	40.6	45.8	49.8	53.4	56.4	59.0	

Table 3

PEARSON PRODUCT-MOMENT CORRELATION COEFFICIENTS, MEANS AND STANDARD DEVIATIONS OF PRE-COURSE AND POST-COURSE ARS FACTOR SCORES OF MATCHED SAMPLE

(N = 1223)

	Pre 1	Pre 2	Pre 3	Pre 4	Post 1	Post 2	Post 3	Post 4
Pre 1 Interest/Affect	-							
Pre 2 Practical Value	-.09							
Pre 3 Negative Affect	.05	.04						
Pre 4 Difficulty	-.05	.04	.05					
Post 1 Interest/Affect	.47	-.04	-.02	.04				
Post 2 Practical Value	.02	.50	-.02	.01	.08			
Post 3 Negative Affect	-.06	.00	.33	.02	-.04	-.07		
Post 4 Difficulty	-.03	.04	-.01	.49	.04	-.04	-.02	
Mean	.040	-.061	.120	-.078	-.041	.063	-.126	.072
Standard Deviation	.895	.972	.894	.9054	1.11	1.09	1.089	1.095

Table 4

SUMMARY OF REGRESSION ANALYSIS OF PRE-AND POST-ARS FACTOR SCORES

DEPENDENT VARIABLE = POST 1

	Multiple R	R^2	R^2 change	Simple R	BETA
Pre 1	.473	.224		.473	.479
Pre 2	.473	.224	.000	-.047	.000
Pre 3	.475	.225	.001	-.018	-.044
Pre 4	.479	.229	.004	.035	.064

Analysis of Variance	DF	Sum of Squares	Mean Square	F	
Regression	4	332.768	83.192		
Residual	1170	1117.096	.955	87.132	p < .001

DEPENDENT VARIABLE = POST 2

	Multiple R	R^2	R^2 change	Simple R	BETA
Pre 1	.015	.000		.015	.065
Pre 2	.507	.257	.257	.503	.521
Pre 3	.514	.264	.007	-.017	-.086
Pre 4	.514	.265	.001	.006	.008

Analysis of Variance	DF	Sum of Squares	Mean Square	F	
Regression	4	327.284	81.821		
Residual	1170	910.038	.778	105.194	p < .001

DEPENDENT VARIABLE = POST 3

	Multiple R	R^2	R^2 change	Simple R	BETA
Pre 1	.061	.004		-.061	-.081
Pre 2	.061	.004	000	.004	-.046
Pre 3	.339	.115	.111	.327	.336
Pre 4	.339	.115	000	.023	.004

Analysis of Variance	DF	Sum of Squares	Mean Square	F	
Regression	4	159.439	39.860		
Residual	117	1231.564	1.053	37.867	p < .001

DEPENDENT VARIABLE = POST 4

	Multiple R	R^2	R^2 change	Simple R	BETA
Pre 1	.033	.001		-.033	-.001
Pre 2	.050	.003	.002	.040	.024
Pre 3	.052	.003	000	-.010	-.038
Pre 4	.494	.244	.241	.492	.493

Analysis of Variance	DF	Sum of Squares	Mean Square	F	
Regression	4	343.322	85.830		
Residual	1170	1065.323	.910	94.264	p < .001

Table 5

MEANS, STANDARD DEVIATIONS, AND INTERFACTOR PRODUCT-MOMENT CORRELATIONS FOR ARS RESIDUALS AND IDEA FACTOR SCORES

(N = 57 classes)

	1	2	3	4	5	6	7	8	9	10	11	12	$\bar{x}$	S.D.
1. Interest/Affect (ARS)	--												.044	.424
2. Practical Value (ARS)	-.12												.046	.329
3. Negative Affect (ARS)	-.03	-.50											-.048	.395
4. Difficulty (ARS)	.27	-.10	.03										.007	.472
5. Organization (IDEA)	-.04	-.18	.38	-.02									-.035	.490
6. Empathy Towards Students (IDEA)	-.18	-.05	.02	.23	-.16								.037	.541
7. Student Cognitive Gain (IDEA)	-.15	-.35	.38	-.14	.03	-.09							-.079	.494
8. Examinations (IDEA)	.07	-.11	-.01	-.25	-.11	.04	.29						-.070	.566
9. Student Personal Gain (IDEA)	-.26	.06	-.02	.25	.00	.31	-.43	-.19					-.023	.488
10. Teaching Style (IDEA)	-.58	-.04	.37	-.32	.06	.07	.02	-.01	.18				-.084	.631
11. Workload (Non-Reading) (IDEA)	-.04	.09	.01	-.82	.06	-.17	.10	.12	-.31	.14			.076	.740
12. Workload (Reading) (IDEA)	.02	-.22	.12	-.01	-.06	-.02	.01	.22	.10	.07	-.19	--	-.108	.672

Table 6 reports the results of the canonical variate analysis between the IDEA factors and the ARS residuals.

Two of the canonical variate pairs reported in Table 6 were significant at the p .001 level. The first canonical R was .897, indicating that the optimal linear combination of IDEA factors and ARS factor residuals accounted for 80 percent of the variance common to the two sets.

The second canonical correlation was .842, accounting for 71 percent of the variance common to the two sets of variables.

Examination of the correlation of each variable with the variates (Table 7) clearly shows that the first canonical variate could be labeled Course Difficulty. The ARS Difficulty residual score and IDEA Student Workload (non-reading) are the variables that correlate the highest with this variate. The reader should note that a positive difficulty residual will result when the observed (i.e., post-course factor) Difficulty factor score represents an easier course than predicted based on the pre-course expectation. The first variate represents a Difficulty factor in the easier direction; the easier the student perceives the course, the higher will be the course variate scores.

The second significant canonical variate pair was termed Positive Interest/Affect. The IDEA factors of Faculty Teaching Style, Student Cognitive Gain, and Course Organization all load significantly with this variate, and all load in the preferred direction. The ARS residuals of Interest/Affect (post-course perception more positive that predicted from pre-course scores) and Negative Affect (post-course perception less negative than predicted from pre-course scores) correlate significantly with this variate.

Table 6

SUMMARY OF CANONICAL VARIATE ANALYSIS FOR ARS RESIDUALS AGAINST IDEA SURVEY FACTORS

(N = 57 classes)

Number of Canonical Variates	Eigenvalue	Canonical Correlation	Wilks Lamda	Chi Square	Degree of Freedom	Significance Level
1	.805	.897	.043	158.474	32	.001
2	.710	.843	.222	75.968	21	.001
3	.188	.434	.766	13.453	12	.500
4	.056	.237	.944	2.931	5	.700

	CANVAR 1	CANVAR 2	CANVAR 3	CANVAR 4
1. Interest/Affect	-.098	.760	-.567	-.439
2. Practical Value	-.016	-.119	.637	-.977
3. Negative Affect	-.236	.598	-.244	-.938
4. Difficulty	1.000	.265	.052	-.093
1. Course Organization	-.052	.307	-.495	-.628
2. Faculty Empathy Toward Students	.127	.216	-.051	.176
3. Student Cognitive Gain	-.072	.597	-.423	.477
4. Examinations (Tests)	-.139	-.199	-.052	.274
5. Student Personal Gain	.016	.240	.269	.487
6. Faculty Teaching Style	-.271	.636	.475	-.186
7. Student Workload (non-reading)	-.855	-.254	.079	.167
8. Student Workload (reading)	-.155	.031	-.479	.272

Table 7

CORRELATION COEFFICIENTS OF VARIABLES WITH
CANONICAL VARIATES
ARS RESIDUAL FACTOR SCORES AND IDEA FACTOR SCORES

Variable	CANVAR I	CANVAR II	CANVAR III	CANVAR IV
ARS 1 Interest/Affect	.184	-.689	-.621	-.324
ARS 2 Practical Value	.010	-.357	.822	-.444
ARS 3 Negative Affect	-.193	.686	-.548	.437
ARS 4 Difficulty	.970	.088	-.177	-.142
IDEA 1 Course Organization	-.121	.328	-.430	-.689
IDEA 2 Faculty Empathy Towards Students	.276	.266	.178	.356
IDEA 3 Student Cognitive Gain	-.228	.413	-.553	.330
IDEA 4 Examinations (Tests)	-.284	-.128	-.277	.475
IDEA 5 Student Personal Gain	.313	.286	.457	.227
IDEA 6 Faculty Teaching Style	.394	.693	.460	-.073
IDEA 7 Student Work Load (Non-Reading)	-.918	-.225	.082	-.051
IDEA 8 Student Work Load (Reading Assignments)	-.038	.090	-.415	.377

A redundancy analysis following the suggestion of Cooley and Lohnes (1976) was completed. The results of the redundancy analysis are reported in Table 8.

Table 8 illustrates that 47 percent of the variance in the ARS residuals is explained by two canonical variate pairs that relate them to the IDEA factors. Conversely, one can also say that 26 percent of the variance in the IDEA factors is explained by two canonical variates which relate the IDEA to the ARS residuals.

In summary, the canonical variate analysis completed to this question yielded two significant canonical variate pairs labeled Difficulty and Positive/Affect. A redundancy analysis illustrated that almost half of the variance in the ARS residual scores can be explained by the pairs of variates which relate the IDEA factors to the ARS residuals.

Summary and Discussion

Two large bodies of literature--one dealing with the theory and measurement of attitudes; the second with student ratings of instruction--have developed along parallel rather than intersecting paths. The literature on the development and formation of attitudes has a strong, though diverse, theoretical base, extending back to 1862 and Spencer's First Principles (Allport, 1967). The literature on student ratings of instruction is not only younger (first recorded use of student ratings of instruction was in 1922) but has developed out of practical rather than theoretical considerations. While social scientists interested in the theory and measurement of attitude press to understand, predict, and, at times, change, human behavior, educators' work on student ratings of instruction has stressed the development of a

Table 8

COMPONENTS OF REDUNDANCY MEASURE
OF ARS RESIDUALS AGAINST
IDEA FACTOR SCORES

Canvar No.	Canonical R	R^2 λ	Variance Extracted	Redundancy
		IDEA		
1	.897	.805	.165	.133
2	.843	.710	.123	.087
3	.434	.188	.150	.028
4	.237	.056	.142	.008
TOTAL			.580	.356
		ARS RESIDUALS		
1	.897	.805	.253	.204
2	.843	.710	.270	.192
3	.434	.188	.348	.065
4	.237	.056	.128	.007
TOTAL			.999	.468

reliable and valid measure of the effectiveness of instructors (usually in order to reward the good and weed out the bad). Current work on student ratings of instruction has begun to stress their use in the improvement of instruction. This study was an attempt to bridge the gap between those two bodies of literature.

The two instruments used in the study were the Adjective Rating Scale (ARS) (Kelly et al., 1976), a measure of course-oriented attitudes, and the Instructional Development and Effectiveness Assessment Survey (IDEA Survey), a student rating of instruction instrument developed at the Center for Faculty Evaluation and Development in Higher Education at Kansas State University. The ARS was completed by 62 classes (1664 students) as both a pre and post-course instrument. The IDEA Survey was given at the end of the semester. Both instruments were factor analyzed. The ARS factor analysis yielded four factors labeled: Interest/Affect, Practical Value, Negative Affect, and Difficulty. The factor analysis of the IDEA Survey resulted in eight factors labeled: Faculty Teaching Style, Faculty Empathy Toward Students, Student Cognitive Gain, Examinations (tests), Student Personal Gain, Course Organization, Student Workload (non-reading), and Student Workload (reading assignments).

Factor scores were generated for each student on each of the factors, and these factor scores became the variables used in the analysis of the question investigated.

The pre-course ARS factor scores were used to predict each of the post-course factor scores and the difference between the predicted post-course factor scores and the factor scores actually observed for each student was

calculated. This difference between post-course predicted scores and post-course observed scores is termed a "residual." Residual scores for each student for each of the ARS factors were calculated, and class means for each of the residuals were computed. The relationship between the ARS residual scores and the IDEA Survey factor scores was then examined.

Residual scores are the difference between the student's actual post-course course-oriented attitude factor scores and the predicted post-course factor scores, based on his pre-course responses. If a student's pre-course attitude has been entirely predictable by his or her pre-course attitude, the residual score would have been zero. There was, however, a significant difference between the student's observed post-course attitudes and the same factors as predicted by the pre-course scores. Hence, the relationship between the ARS residualized scores and the IDEA factor scores was examined through canonical variate analysis. Table 9 summarizes the dimensions that explained the relationship between the ARS residualized factor scores and the IDEA factor scores. Two dimensions (Course Difficulty and Positive Interest/Affect) are significantly related to residualized attitude toward a course as measured by the ARS.

The primary focus of this analysis was to examine the direction and intensity of the relationship between student ratings of instruction and student attitudes toward a course. This time student attitudes were represented by those post-course attitudes unaccounted for by pre-course expectations. The direction of this overlap is illustrated in Table 9. However, the strength of this overlap needs to be understood.

Table 9

DIMENSIONS THAT DESCRIBE THE OVERLAP BETWEEN STUDENTS' COURSE-ORIENTED ATTITUDE RESIDUAL SCORES AND STUDENT RATINGS OF INSTRUCTION

Dimension 1 Course Difficulty*	Dimension 2 Positive Interest/Affect
Course is Easier than predicted (ARS)**	More Positive Interest/Affect than predicted (ARS)
Low Student Workload (non-reading) (IDEA)	Less Negative Affect than predicted (ARS)
Positve Faculty Teaching Style (IDEA)	Moderately More Practical Value than predicted (ARS)
Moderately Positive Student Personal Gain (IDEA)	Very Positive Faculty Teaching Style (IDEA)
	Moderately High Student Cognitive Gain (IDEA)
	Moderately Good Course Organization (IDEA)

* Only significant canonical functions ($p < .001$) are included.

** Only factors or residuals that correlate $\geq .30$ with variate are included.

First, the amount of variance in post-course attitudes, unaccounted for by pre-course attitudes, must be examined. This was accomplished by subtracting the squared multiple correlation of each post-course ARS factor and the pre-course factors from one. The squared multiple correlation (R^2) is:

$$R^2 = \frac{\text{The variation in Y explained by the linear influence of the independent variables}}{\text{The total variation in Y}} .$$

"Y" represents the post-course ARS factor scores, and the independent variables are the pre-course ARS factor scores. The difference between R^2 and 1 is the amount of variation in the post-course factors that is not accounted for by the influence of the pre-course factors. The results are as follows: ARS1 (Interest/Affect) = .77, ARS2 (Practical Value) = .74, ARS3 (Negative Affect) = .89, and ARS4 (Difficulty) = .76. Each of these figures represents the percent of post-course attitude variance unaccounted for by pre-course expectations. These figures are then compared with the redundancy analysis calculated for the ARS residualized factor scores and the IDEA factor scores. The results of the redundancy analysis showed that 47 percent of the variance in the ARS residuals is explained by the variance in the IDEA scores; or to be more precise, approximately 47 percent of the 79 percent (average variance in the post-variance of the ARS residualized factor scores can be accounted for by students' responses on the IDEA. To be able to explain over 37 percent of the variance left-over after the pre-course expectations are taken into account implies that the relationship between course-oriented attitudes and ratings of instruction is measurable even after students' course-oriented expectations have been controlled. The data collected in this study would support the conclusion that the IDEA Survey produces ratings that will explain

a moderate amount of course-oriented attitude unaccounted for by pre-course expectations.

Implications of the Findings

Several implications seem reasonable to draw from the results of this investigation. First, and perhaps most important, it is clear that students' course-oriented attitudes and student ratings of instruction have much in common, perhaps enough in common to be considered the same construct.

Rokeach (1970) offered one explanation of why research on attitude measurement and the rapidly expanding literature on student ratings of instruction have developed along parallel rather than intersecting paths. He noted that attitude theorists are generally more interested in the theory and measurement of attitudes toward situations, across objects. "The splitting-off of attitude-toward-situation from attitude-toward-object has, (in his opinion) severely retarded the growth of attitude theory" (p. 119). If a course can be considered a situation (as this investigator believes), then this study will hopefully encourage social scientists who are interested in the theory and measurement of attitudes to include courses, or a course, as an acceptable attitude generating situation. By the same token, it is also hoped that researchers interested in student ratings of instruction will begin to draw more from the theories and techniques of attitude measurement in developing their instruments, designing their research studies, and interpreting their findings.

Since student ratings of instruction have become a part of the evaluation of instruction in many colleges, it is imperative that we begin to develop theoretical or conceptual models that will help us interpret such rating.

Current theories of attitude formation and measurement seem to be a useful place to search for our first tentative models. Such a search is more difficult than it might appear, first, because, as already mentioned, most of the research on attitudes has concentrated on attitudes toward objects across situations and, second, because the literature on attitude formation and measurement is itself so diverse. Early in this report, it was stated that Rokeach's definition of attitude would be accepted as the "working" definition for this study. No definition of student ratings of instruction was given because no satisfactory definition could be found in the literature. Rokeach gave the following extended definition of attitude.

> An attitude is a relatively enduring organization of interrelated beliefs that describe, evaluate, and advocate action with respect to an object or situation, with each belief having cognitive, affective, and behavioral components. Each of these beliefs is a predisposition that, when suitably activated, results in some preferential response toward the attitude object or situation, or toward others who take a position with respect to the attitude object or situation, or toward the maintenance or preservation of the attitude itself. Since an attitude object must always be encountered within some situation about which we also have an attitude, a minimum condition for social behavior is the activation of at least two interacting attitudes, one concerning the attitude object and the other concerning the situation, (p. 132).

This investigator believes that the items on a typical student rating of instruction form fit the definition of a set of interrelated beliefs mentioned above, and that the total set of these beliefs forms an attitude toward a course. Hence, student ratings of instruction are, in essence, measures of attitude. It is not being proposed that the individual items on a student rating form measure attitudes, but rather that they measure beliefs; some of these beliefs are descriptive of class or instructor characteristics; others are evaluative. None of the items on student rating

forms measure behavior, although some rating forms call for students to express a behavioral preference such as "I would like to take another class from this instructor."

There are many difficulties that occur in interpreting a definition of attitude even when one is accepted and to say that items on a student rating of instruction form represent a set of beliefs that comprises on attitude does not mean that definitional problems are solved. First, we do not know if student rating forms tap all the beliefs that constitute an attitude toward a course, nor do we know which of these beliefs are subject to change based on the course content, teacher behaviors, interaction with other students, etc. Nor, is it clear, even if we accept Rokeach's definition, exactly what is the object and what is the situation to which he refers. Is the instructor the object and the course the situation, or are the instructor and the course both objects with the educational environment the situation? Such questions still have to be answered.

Placing student ratings of instruction into a theoretical framework is more than just an intellectual game. Practical considerations arise when we consider the behavioral components of attitudes. "Does a student with a positive attitude toward a course study harder, learn more, take another course from the instructor, take another course in the same area, change majors, stay in school?" are all relevant questions in higher education today. Without a theoretical framework, investigators have relied on correlational analysis to attempt to answer such questions.

A second implication that can be drawn from this study is that students' post-course attitudes are not completely predictable from their pre-course

attitudes. In fact, far more of the post-course variance in attitudes can not be predicted by pre-course attitudes than can be predicted; and this unpredictable variance in post-course attitude is related to student ratings of instruction. For example, a course that is perceived as easier than expected, based on pre-course expectations, is positively related not only to non-reading workload (as might be expected) but also to such items as "Instructor spoke with expressiveness and variety in tone of voice," "Instructor seemed enthusiastic about subject matter," and "Instructor made presentations which were not dry and dull." By the same token, courses which received more positive ARS Interest/Affect scores than predicted and lower Negative Affect scores than predicted also received high scores on Faculty Teaching Style (items mentioned above) as well as above average ratings on students' perceptions of their Cognitive Gain. In other words, students' post-course attitudes that differ significantly from what would be predicted based on their pre-course responses are related not only to course variables such as Difficulty or Amount of Work but are also related to the students' perceptions of instructor behaviors.

Limitations of the Study

The implications of this study are limited by two classes of variables, instrumentation, and environmental factors including sampling.

Low or unknown reliability is a pervasive problem with measures of attitude generally, and this study has not escaped this issue. Little information was available on either the test-retest reliability or the parallel form reliability of the measures used in this study.

Far more information was available concerning the Adjective Rating Scale than was available for the IDEA Survey. Both of these instruments are appealing to faculty for various reasons. The IDEA Survey has a unique method of reporting student opinion data back to faculty, a method that takes instructors' goals for the course as well as student perceptions into account. The IDEA Survey is currently in use in over 30,000 classrooms across the county, but empirical data on the instrument is sorely lacking.

Although more information is available on the ARS, further investigation of this instrument--its validity and reliability--and its factor structure when taken over other levels of programming appears to be a promising direction for further research. The ARS is easy to administer, and, with its course, as opposed to teacher behavior orientation, is relatively non-threatening to most faculty.

This study was also limited by environmental factors. The faculty participating in the study were self-selected. Although an attempt was made to include all departments in the college and both lower level and upper level courses, the only incentive for faculty to participate in the study was interest in their students' opinions and friendship with the principal investigator.

There may be unknown problems with the student sample. As is demonstrated by the drop in numbers from the pre-course survey to the post-course survey, to the matched sample, students will drop out of courses, will not show up on the day an instrument is given, and will sometime incorrectly include their student numbers (or fail to include them at all) on an instrument. It is impossible to determine whether students who dropped a course had

substantially different attitudes toward the course than did students who did not drop the course.

Finally, there is a limitation common to all multivariate analyses. Canonical variate analysis allows the investigator to maintain the integrity of distinct variable sets while examining the relationship between the two sets. It was an appropriate statistical procedure for this study. However, the results, as with most multivariate procedures, are difficult to interpret. Cross validation of the results in the manner suggested by Wood and Erskin (undated) is also needed.

Recommendations for Future Research

Several areas of further research have already been mentioned in this chapter. Additional analysis of both the ARS and the IDEA Survey was recommended, as was additional research into various types of student rating forms depending on the potential use of the data. In addition, far more research (and thinking) must be done on the relationship between student attitudes toward a course and student ratings of a course. It is possible that the wording of the items on the student rating form is a critical factor in determining whether or not student ratings tap the same dimension as student attitudes toward a course. Perhaps items that only ask students to be observers of course or teacher characteristics are different from those that ask students how they feel about those same courses or teacher characteristics.

Additional work needs to be done on identifying interactions that may help us explain the relationship of attitude and ratings to other subjects, objects, and context variables. Studies, for example, which seek possible

interactions between different course designs, course attitudes, and course achievement need to be continued. Similarly the relationship of student sex to course-oriented attitude and student ratings in unclear. Separate factor analytic studies describing the structures of response for males and females with respect to course attitude would be productive as would studies which attempt to relate course-oriented attitude to actual behaviors.

References

Allport, G. Attitudes. Attitude theory and measurement. New York: John Wiley and Sons, 1967.

Centra, J. A. Evaluating college teaching: The rhetoric and the research. Current Issues in Higher Education, 1972, 27 (Annual), 225-233.

Cooley, W. W., & Lohnes, P. R. Evaluation research in education: Theory, principle and practice. New York: John Wiley and Sons, 1976.

Costin, F., Greenough, F., & Menges, R. Student ratings of college teaching: Reliability, validity, and usefulness. Review of Educational Research, 1971, 41, 511-535.

Follman, J. Student ratings of faculty teaching effectiveness: Rater or ratee characteristics. Research in Higher Education, 1975, 3, 155-167.

French-Lazovik, G. Predictability of students' evaluations of college teachers from component ratings. Journal of Educational Psychology, 1972, 66 (3), 373-385.

Hogan, T. P. Similarity of student ratings across instructors, courses and time. Research in Higher Education, 1973, 1, 149-154.

Holzemer, W. L. Student ratings of instructional effectiveness and their relations to the college classroom environment. Unpublished doctoral dissertation, Syracuse University, 1975.

Hoyt, D. Instructional Effectiveness: 1. measurement of effectiveness. Kansas State University, Office of Educational Research, Research Report #6, 1969.

Kelly, E. F. A study of student attitudes in Freshman English. Syracuse University: Center for Instructional Development, 1975.

Kelly, E. F., Pascarella, E. T., Terenzini, P. T., & Chapman, D. The development and use of the Adjective Rating Scale: A measure of attitudes toward courses and programs. Syracuse University: Center for Instructional Development, 1976.

Kulick, J. A. & Kulik, C. C. Student ratings of instruction. Teaching of Psychology, December 1974, 1 (2), 51-57.

McKee, B. G. The Relationship Between Students' Course-Oriented Attitudes and Student Ratings of Instruction: A Canonical Variate Analysis. Rochester, N.Y.: NTID (ERIC Document Reproduction Service No. ED 156 698).

Menges, R. J. The new reporters: Students rate instruction. In J. B. Lon Hefferlin (Quarterly source book ed.), Evaluating Teaching and Learning (Issue Editor C. R. Pace). San Francisco: Jossey-Bass, 1973, 4, 59-75.

Naftulin, D., Ware, J., & Donnelly, F. The Doctor Fox lecture: A paradigm of educational seduction. Journal of Medical Education, 1973, 48, 630-635.

Rokeach, D. Beliefs, attitudes and values. San Francisco: Jossey-Bass, Inc., 1970.

Rosenshine, B. Teacher behavior and student attitudes revisited. Journal of Educational Psychology, 1973, 63 (2), 177-180.

Scott, D. L. The evaluation of college teaching. Presentation at the Center for Instructional Development, Syracuse University, October, 1976, Syracuse, New York.

Tatsuoka, M. M. Multivariate analysis. New York: John Wiley and Sons, 1971.

Wood, D. A., & Erskine, J. A. Analytical strategies in canonical correlation: A review and application. Undated manuscript, Graduate School of Business, Indiana University.